Retrieval Enhancements for Task-Based Web Search

The Oral Exam of
Michael Völske

To Obtain the Academic Degree of
Dr. rer. nat.

Web Technology & Information Systems Group
Bauhaus-Universität Weimar

www.uni-weimar.de www.webis.de

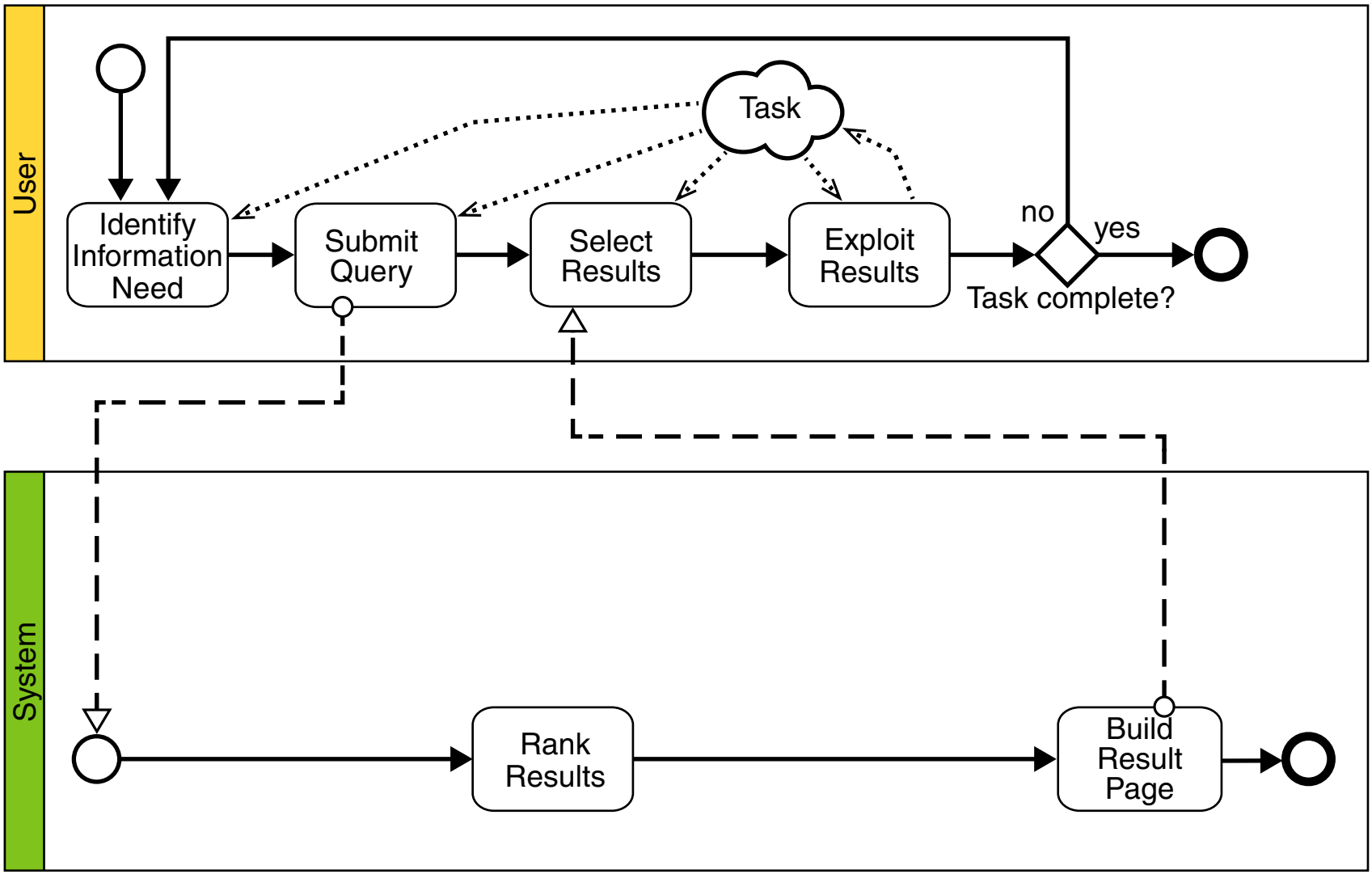
Retrieval Enhancements for Task-Based Web Search

Outline

1. Task-based Search
2. **Contribution 1:** Writing Tasks
3. **Contribution 2:** Axiomatic Result Reranking
4. Summary

Retrieval Enhancements for Task-Based Web Search

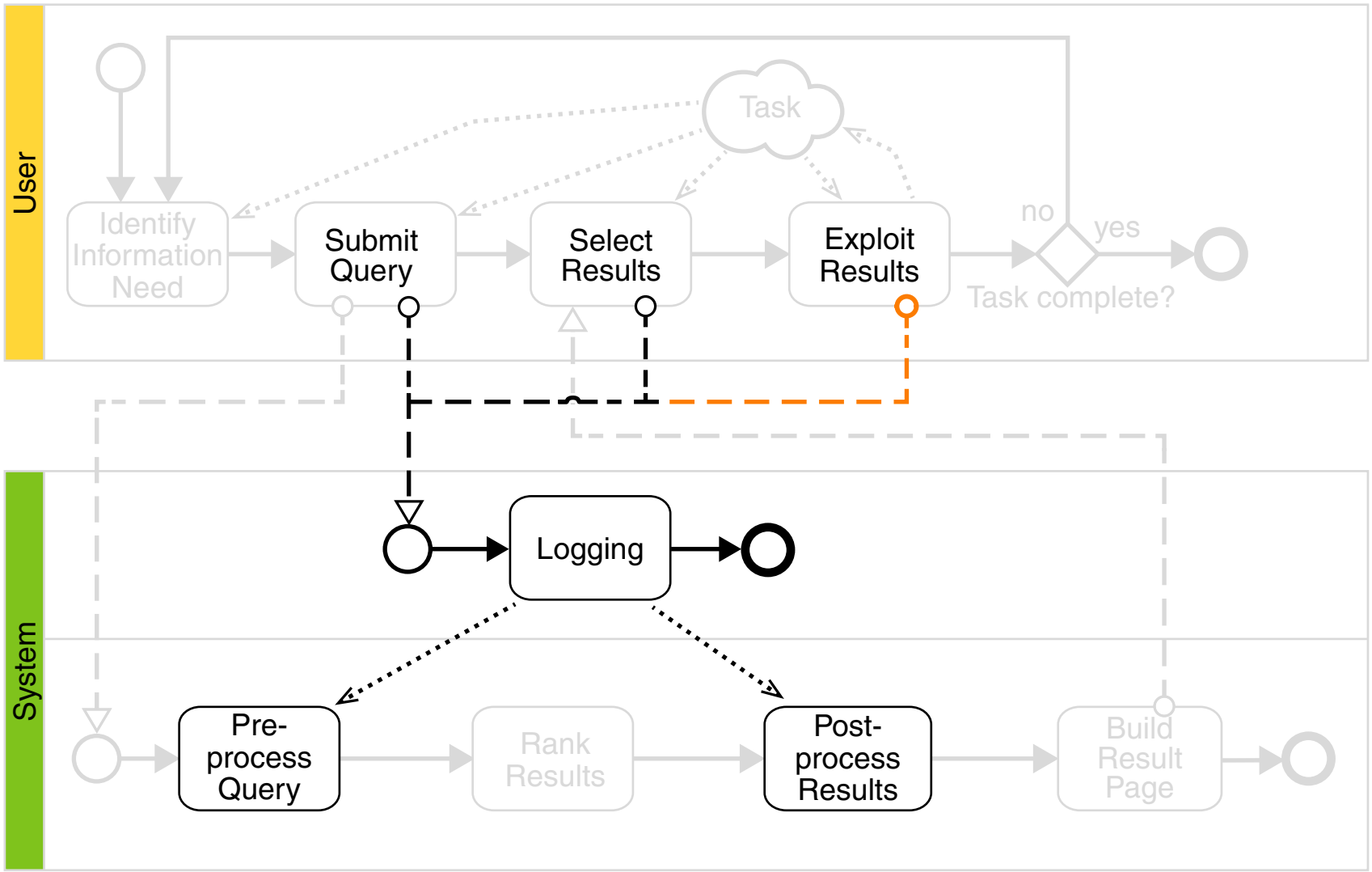
The Task-Based Web Search Process



© Völske 2019

Retrieval Enhancements for Task-Based Web Search

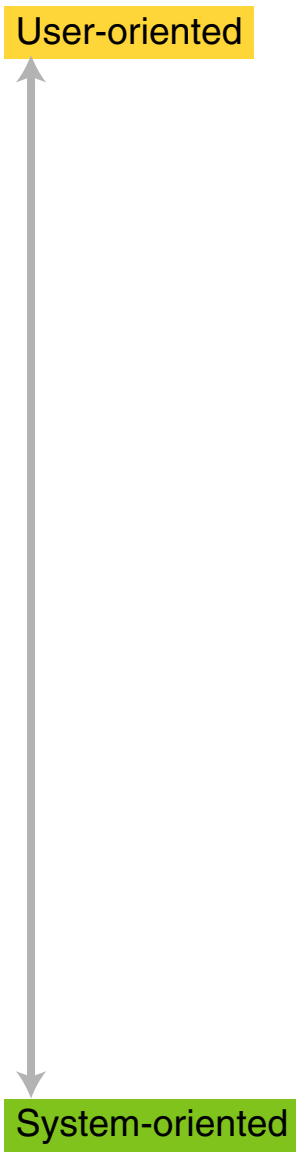
The Task-Based Web Search Process



© Völske 2019

Retrieval Enhancements for Task-Based Web Search

Contributions



© Völske 2019

Retrieval Enhancements for Task-Based Web Search

Contributions

User-oriented

Understanding and Supporting Writing Tasks
[Potthast, Hagen, Völske, Stein; ACL 2013]
[Potthast, Hagen, Völske, Stein; EuroHCIR 2013]
[Hagen, Potthast, Völske, Stein; CHIIR 2016]
[Vakkari, Völske, Hagen, Potthast, Stein; TPDL 2018]
[Vakkari, Völske, Potthast, Hagen, Stein; IP&M 2019]

Result Presentation for Exploratory Search
[Gollub, Völske, Hagen, Stein; JCDL 2014]
[Völske, Gollub, Hagen, Stein; WOSP 2014]

Categorizing Question Queries
[Völske, Braslavski, Hagen, Lezina, Stein; CIKM 2015]

Query-Task Mapping
[Völske, Fatehifar, Stein, Hagen; SIGIR 2019]

Automatic Summarization
[Völske, Potthast, Syed, Stein; NewSum 2017]
[Syed, Völske, Potthast, Lipka, Stein, Schütze; INLG 2018]

Axiomatic Result Reranking
[Hagen, Völske, Göring, Stein; CIKM 2016]

System-oriented

© Völske 2019

Retrieval Enhancements for Task-Based Web Search

Contributions

User-oriented

Understanding and Supporting Writing Tasks
[Potthast, Hagen, Völske, Stein; ACL 2013]
[Potthast, Hagen, Völske, Stein; EuroHCIR 2013]
[Hagen, Potthast, Völske, Stein; CHIIR 2016]
[Vakkari, Völske, Hagen, Potthast, Stein; TPDL 2018]
[Vakkari, Völske, Potthast, Hagen, Stein; IP&M 2019]

Result Presentation for Exploratory Search
[Gollub, Völske, Hagen, Stein; JCDL 2014]
[Völske, Gollub, Hagen, Stein; WOSP 2014]

Categorizing Question Queries
[Völske, Braslavski, Hagen, Lezina, Stein; CIKM 2015]

Query-Task Mapping
[Völske, Fatehifar, Stein, Hagen; SIGIR 2019]

Automatic Summarization
[Völske, Potthast, Syed, Stein; NewSum 2017]
[Syed, Völske, Potthast, Lipka, Stein, Schütze; INLG 2018]

Axiomatic Result Reranking
[Hagen, Völske, Göring, Stein; CIKM 2016]

System-oriented

© Völske 2019

Contribution 1:
Understanding and Supporting Writing Tasks

Understanding and Supporting Writing Tasks

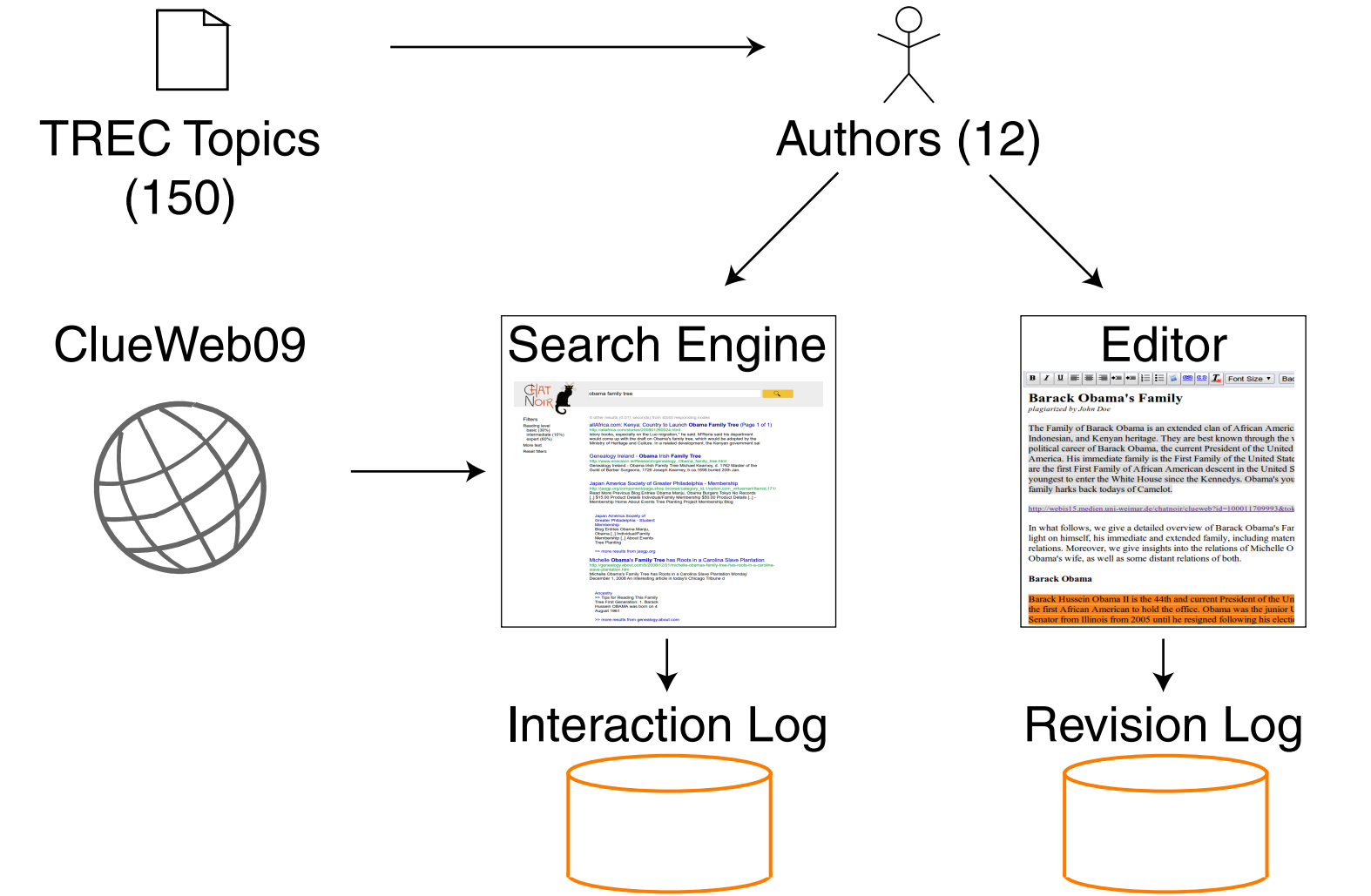
Motivation

1. How do authors interact with search engines during complex writing tasks?

2. What are implications for search engine design?

©Völske 2019

The Webis-TRC-12 Dataset

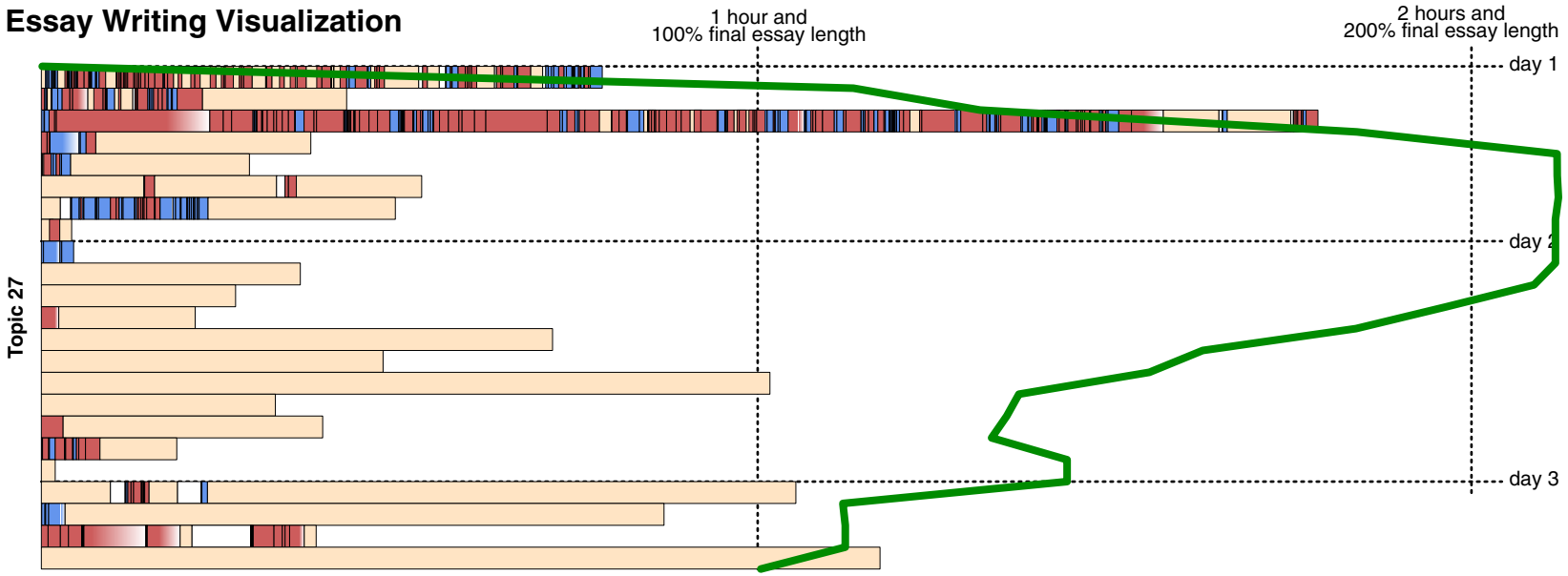


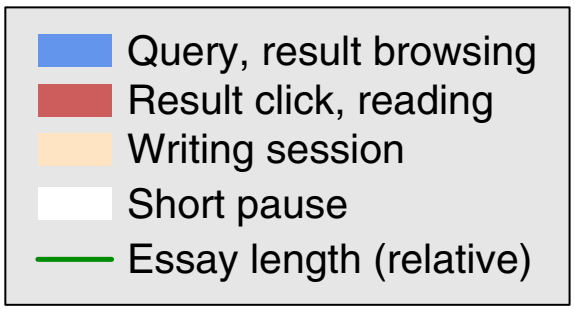
©Völske 2019

Writing and Search Behavior

© Völske 2019

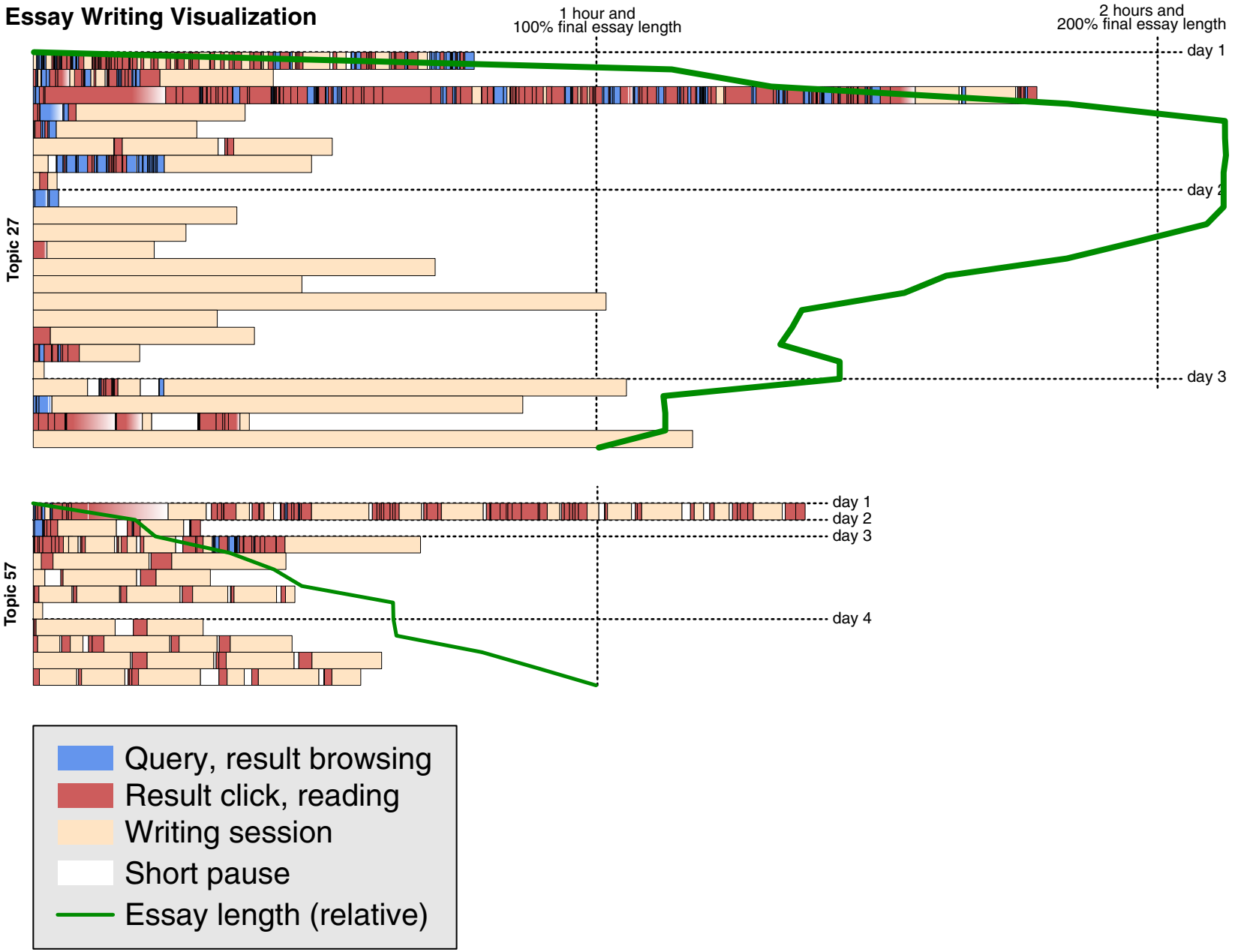
Essay Writing Visualization





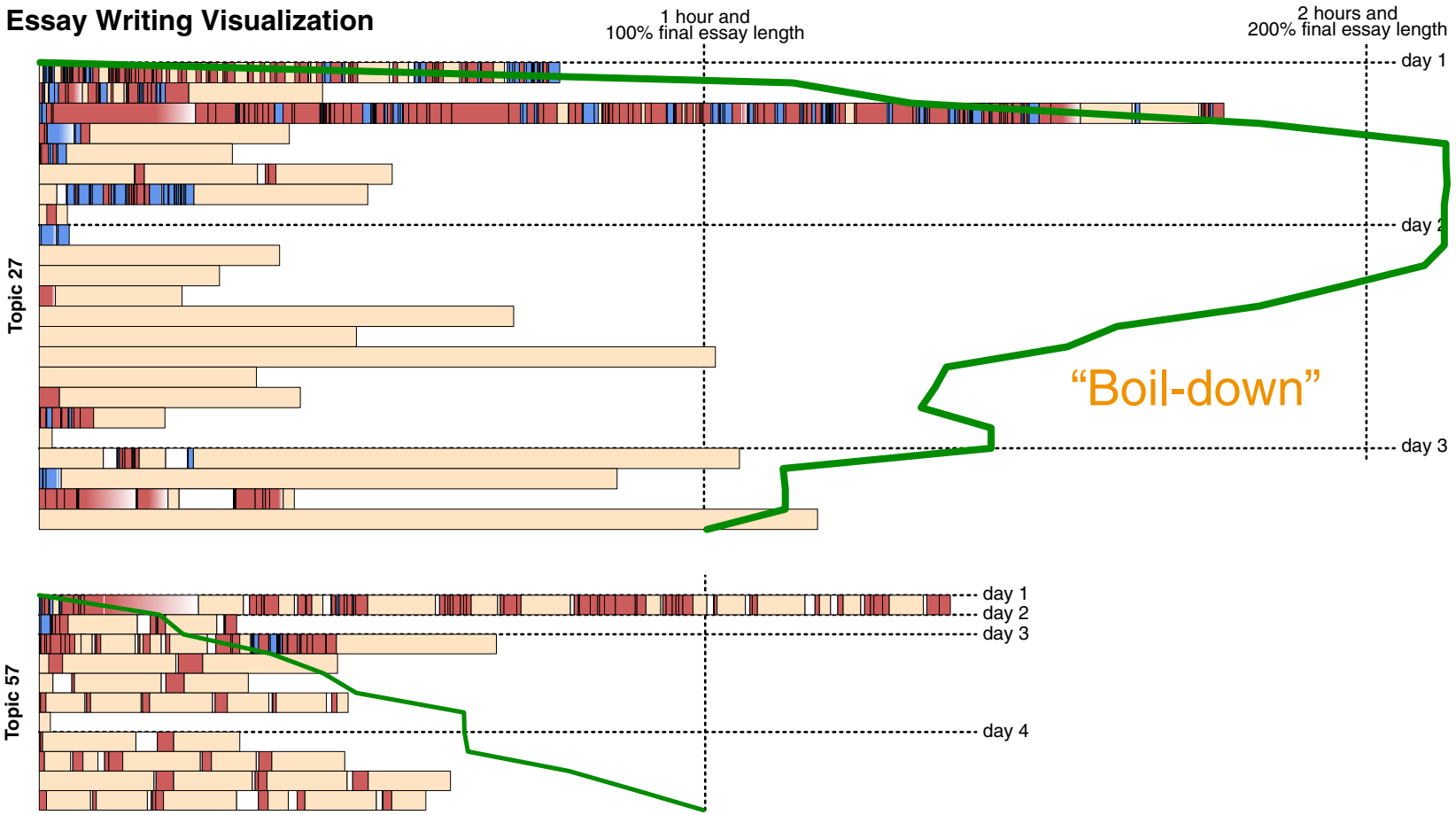
©Völske 2019

Essay Writing Visualization



© Völske 2019

Essay Writing Visualization



Query, result browsing
Result click, reading
Writing session
Short pause
Essay length (relative)

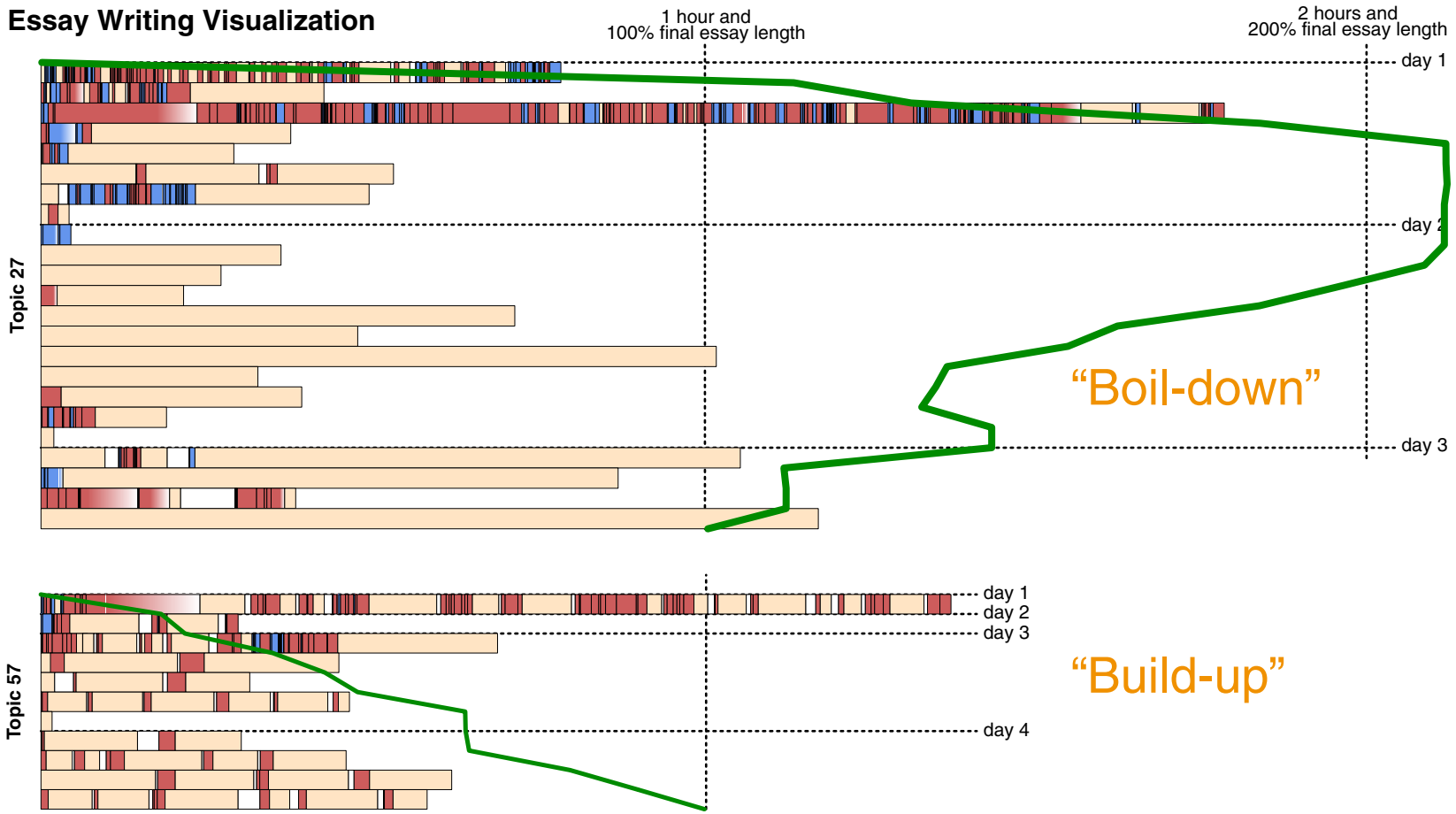
Boil-down Strategy

Two working phases:

1. Gather Material
2. Write text

©Völske 2019

Essay Writing Visualization



- Query, result browsing
- Result click, reading
- Writing session
- Short pause
- Essay length (relative)

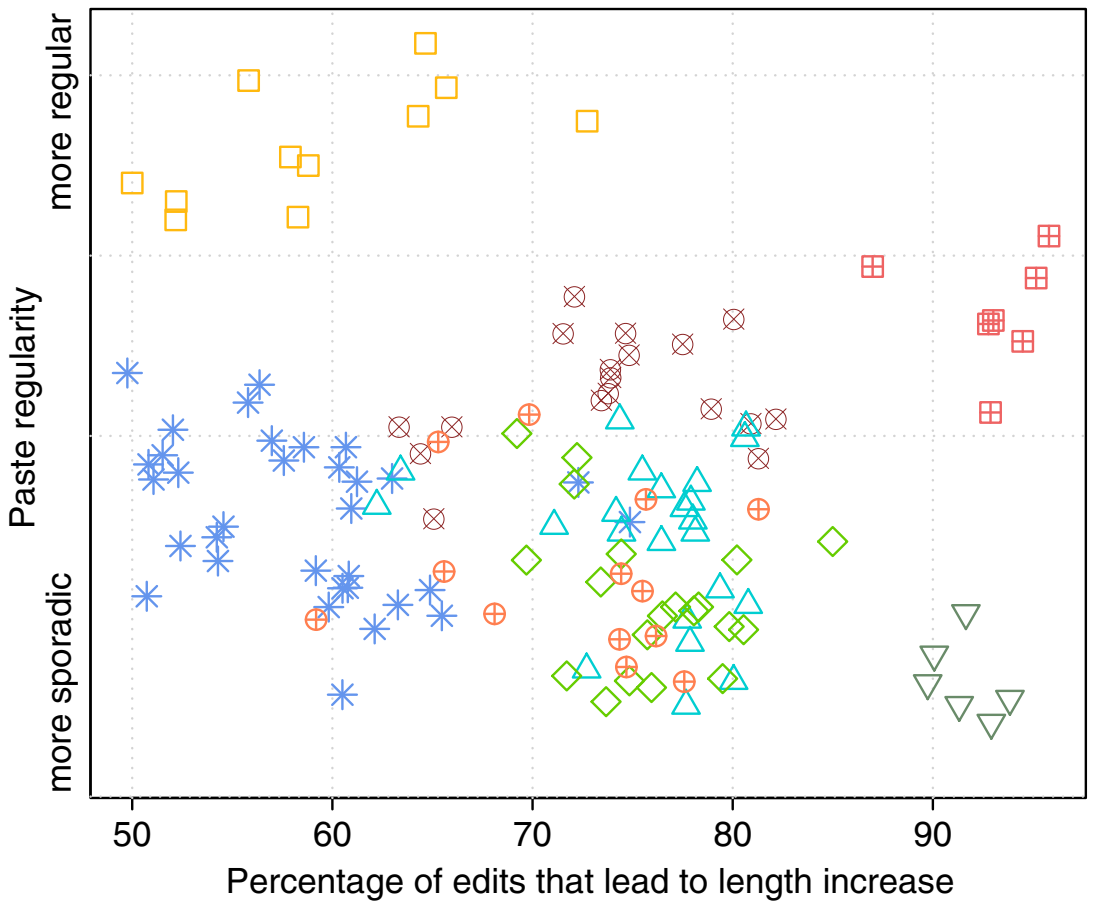
Boil-down Strategy
Two working phases:
1. Gather Material
2. Write text

Build-Up Strategy
1. Gather Source
2. Integrate
3. Repeat

©Völske 2019

Writing and Search Behavior

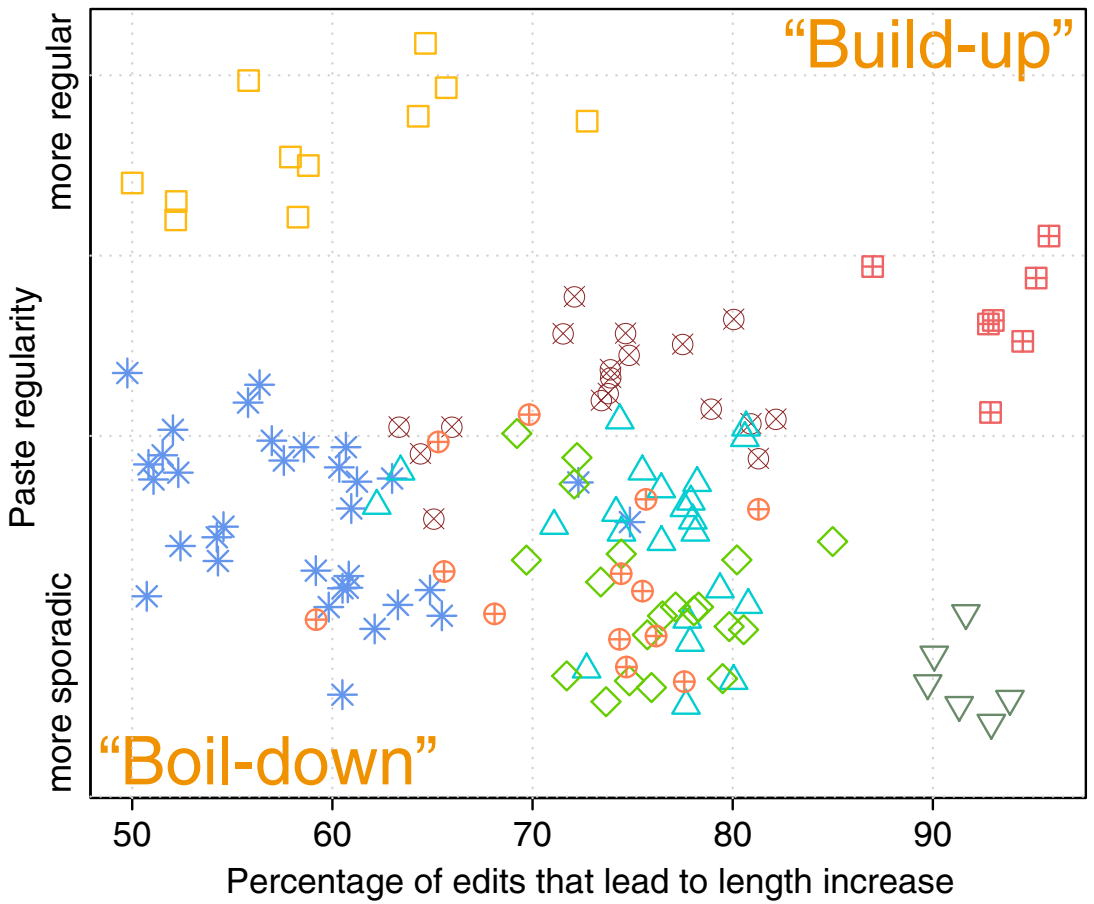
Authors are Quite Consistent!



©Völske 2019

Writing and Search Behavior

Authors are Quite Consistent!



© Völske 2019

Adapting Search to the Author

Towards Predicting Search Result Usefulness

Objective: Search engine should adapt to the author's strategy.

Can we predict which documents will be useful?

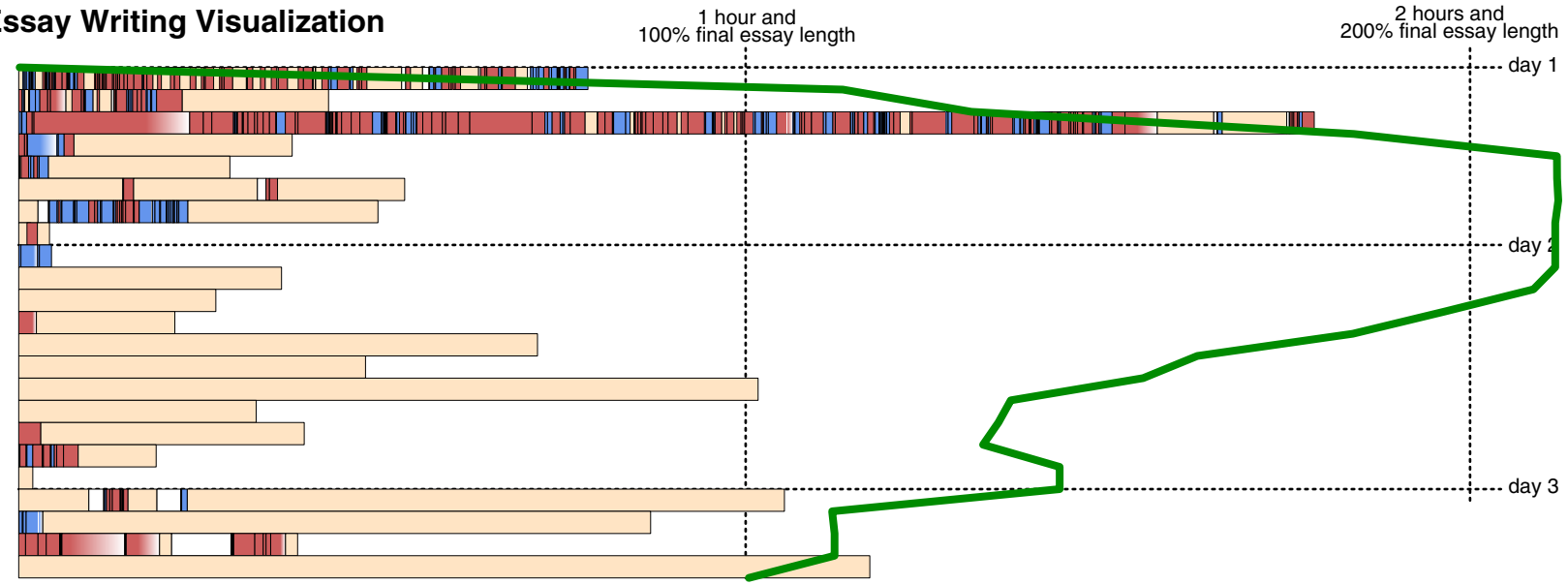
Can we predict authors' degree of retrieval success?

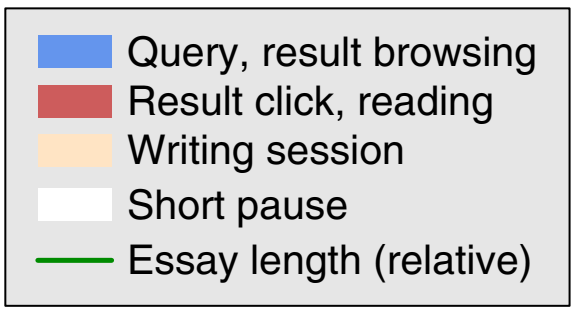
© Völske 2019

What Do I Mean By Retrieval Success

© Völske 2019

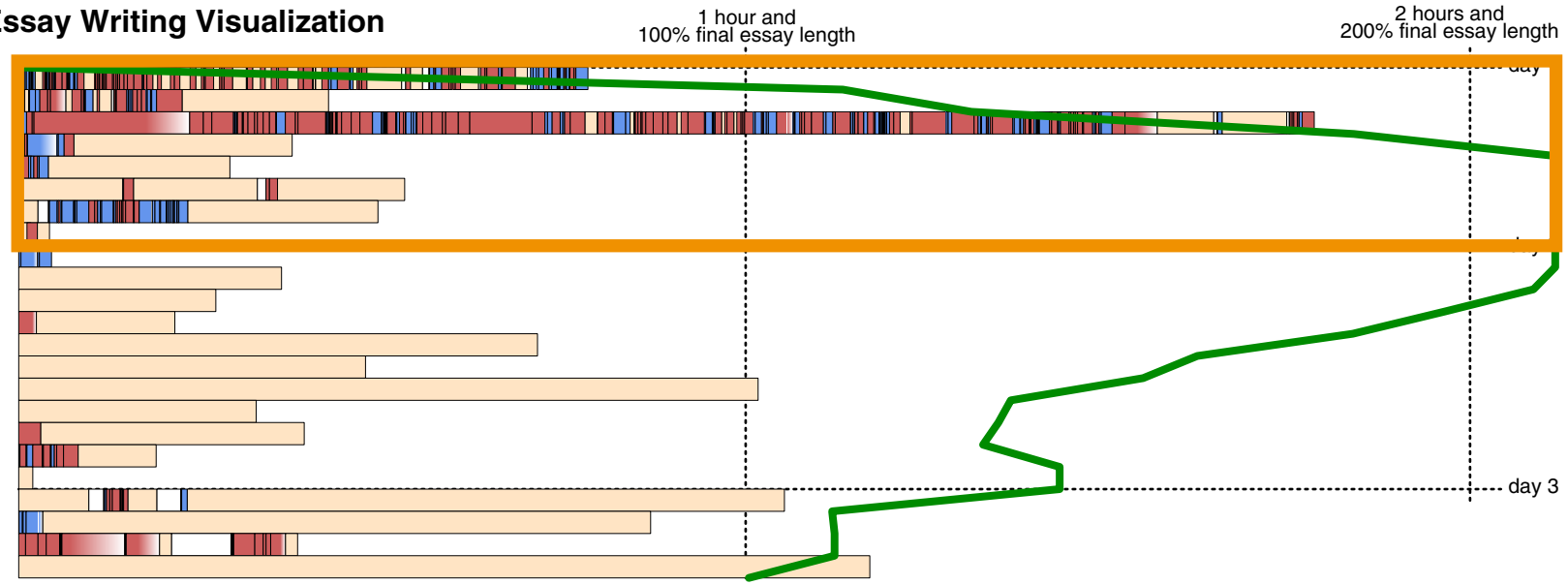
Essay Writing Visualization





©Völske 2019

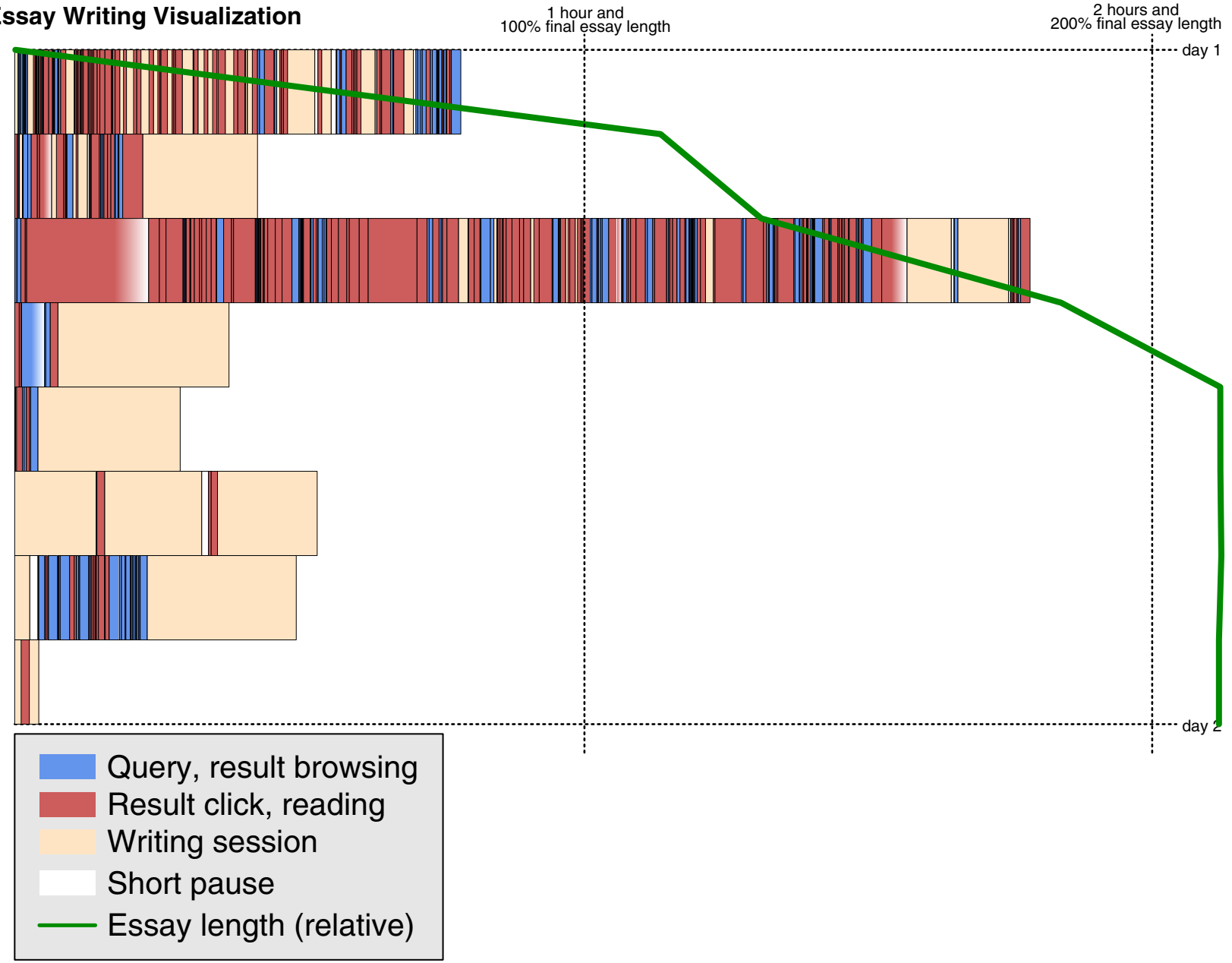
Essay Writing Visualization



Query, result browsing
Result click, reading
Writing session
Short pause
Essay length (relative)

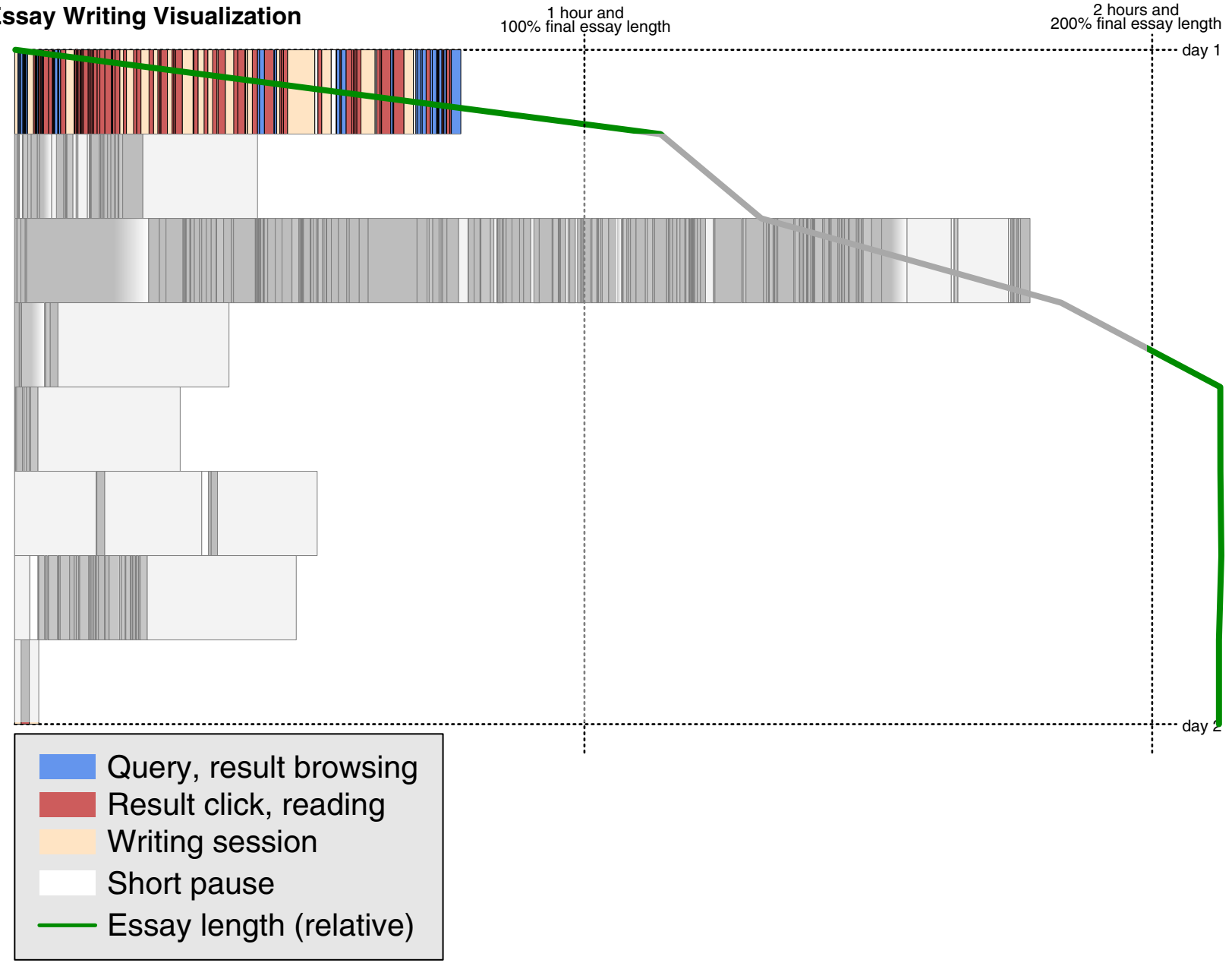
©Völske 2019

Essay Writing Visualization



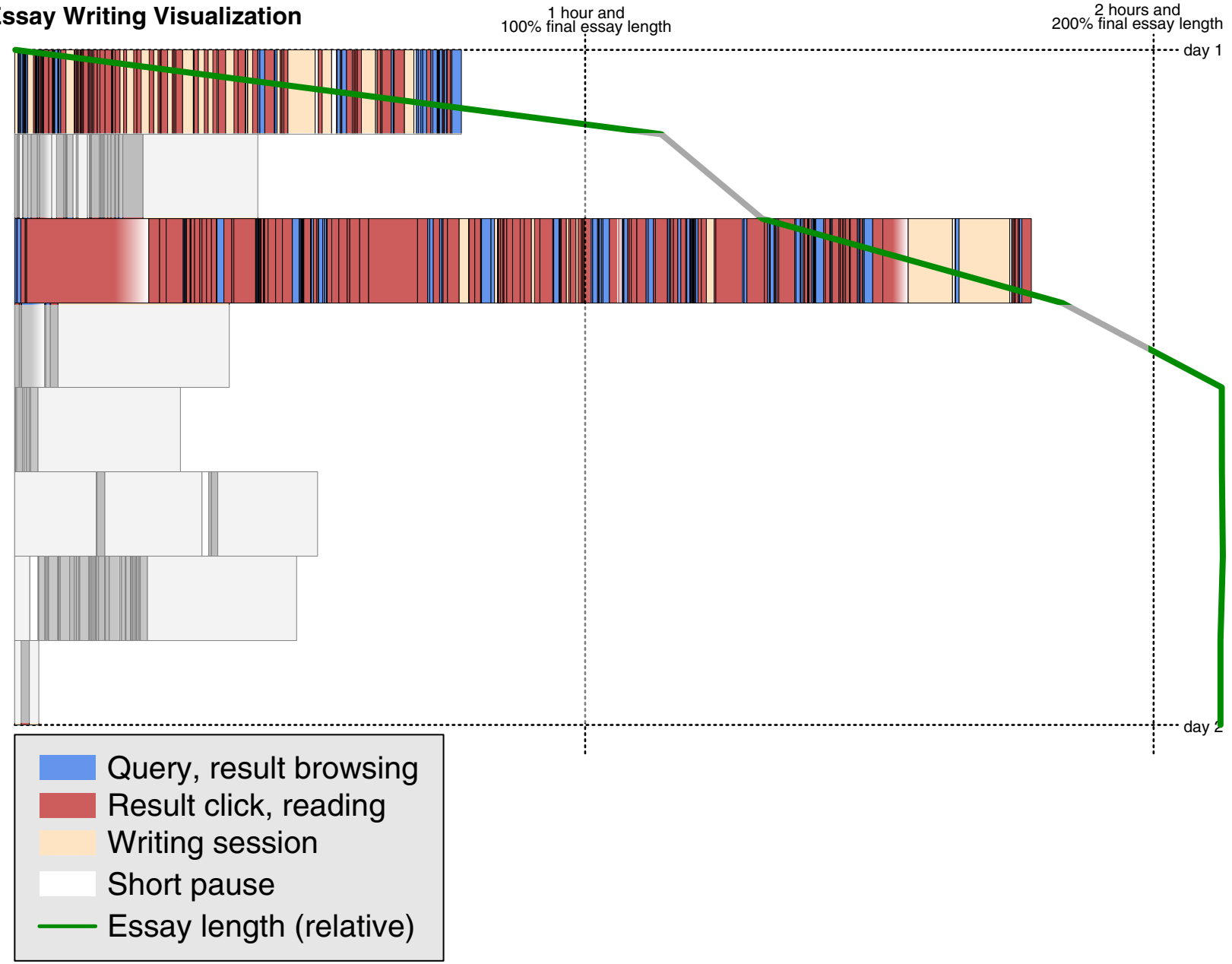
© Völske 2019

Essay Writing Visualization



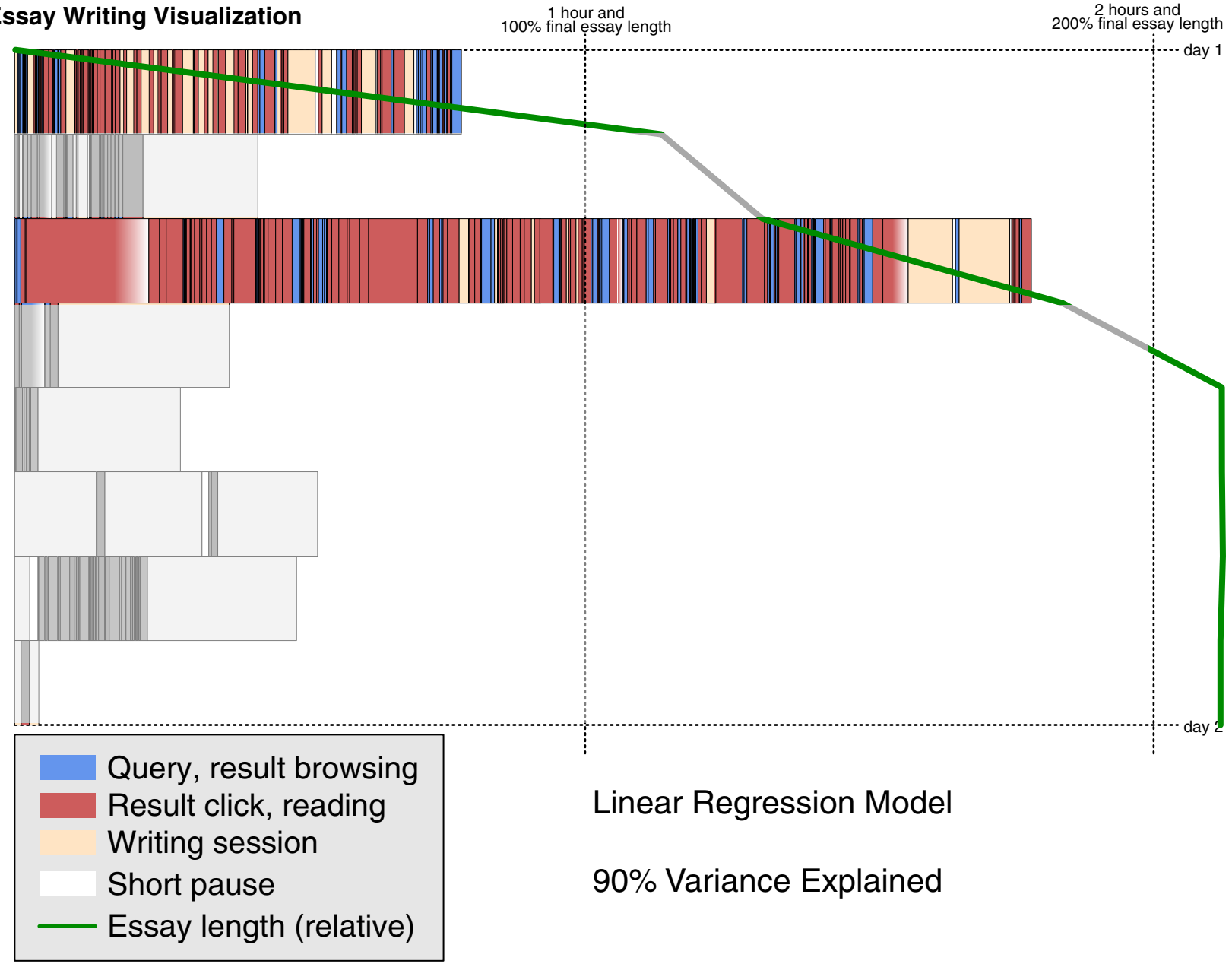
© Völske 2019

Essay Writing Visualization



©Völske 2019

Essay Writing Visualization



©Völske 2019

Contribution 2:
Enhancing Result Rankings with Axioms

Enhancing Result Rankings with Axioms

Four in hand | Tie-a-Tie.net
https://www.tie-a-tie.net/fourinhand/
Looking for an easy and versatile **knot**? Then the "**Four in Hand**" is the one. It is one of the simplest **tie knots**, making it a perfect pick for men new to wearing **ties**.

How to Tie a Tie - Apps on Google Play
https://play.google.com/store/apps/details?id=com.artelplus.howtotie&hl=en_US
Download this app and don't waste your time and money searching the internet and trying to find out **how to tie** a decent knot on your own. Here they are - tens ...

How to Tie a Half Windsor Knot | The Tie Bar
https://www.thetiebar.com/how-to-tie/half-windsor
Learn **how to tie** a Half Windsor Knot and quickly become a Half Windsor expert with our easy, step-by-step instructions.

How To Tie A Tie Knot - 17 Different Ways of Tying Necktie Knots
https://www.realmenrealstyle.com/how-to-tie-a-tie/
I'm showing you 18 ways to **tie a necktie**, so you will be able to tie all the best tie knots – starting with the quickest and easiest tie knot, the Simple or Oriental ...

How to Tie a Tie: The Complete Guide | The Art of Manliness
https://www.artofmanliness.com/articles/how-to-tie-a-tie/
May 10, 2019 - It's a sad fact, but there are grown men who don't know **how to tie a tie**. If they have a big interview that afternoon, they'll go shopping for a ...

How To Tie A Windsor Knot | Ties.com
https://www.ties.com/how-to-tie-a-tie/windsor
Although the Duke of Windsor never specifically used the Windsor knot, he did favor a wide triangular knot. In actuality, the Duke achieved his trendsetting look by tying a Four-in-Hand with specially made wide and extra thick **ties**. The Windsor knot was invented by the public as a ...
How To Tie A Simple Knot · Half Windsor · "Double Windsor" knot

© Völske 2019

Enhancing Result Rankings with Axioms

Four in hand | Tie-a-Tie.net
https://www.tie-a-tie.net/fourinhand/
Looking for an easy and versatile **knot**? Then the "**Four in Hand**" is the one. It is one of the simplest **tie knots**, making it a perfect pick for men new to wearing **ties**.

How to Tie a Tie - Apps on Google Play
https://play.google.com/store/apps/details?id=com.artelplus.howtotie&hl=en_US
Download this app and don't waste your time and money searching the internet and trying to find out **how to tie** a decent knot on your own. Here they are - tens ...

How to Tie a Half Windsor Knot | The Tie Bar
https://www.thetiebar.com/how-to-tie/half-windsor
Learn **how to tie** a Half Windsor Knot and quickly become a Half Windsor expert with our easy, step-by-step instructions.

How To Tie A Tie Knot - 17 Different Ways of Tying Necktie Knots
https://www.realmenrealstyle.com/how-to-tie-a-tie/
I'm showing you 18 ways to **tie a necktie**, so you will be able to tie all the best tie knots – starting with the quickest and easiest tie knot, the Simple or Oriental ...

How to Tie a Tie: The Complete Guide | The Art of Manliness
https://www.artofmanliness.com/articles/how-to-tie-a-tie/
May 10, 2019 - It's a sad fact, but there are grown men who don't know **how to tie a tie**. If they have a big interview that afternoon, they'll go shopping for a ...

How To Tie A Windsor Knot | Ties.com
https://www.ties.com/how-to-tie-a-tie/windsor
Although the Duke of Windsor never specifically used the Windsor knot, he did favor a wide triangular knot. In actuality, the Duke achieved his trendsetting look by tying a Four-in-Hand with specially made wide and extra thick **ties**. The Windsor knot was invented by the public as a ...
How To Tie A Simple Knot · Half Windsor · "Double Windsor" knot

How To Tie A Tie Knot - 17 Different Ways of Tying Necktie Knots
https://www.realmenrealstyle.com/how-to-tie-a-tie/
I'm showing you 18 ways to **tie a necktie**, so you will be able to tie all the best tie knots – starting with the quickest and easiest tie knot, the Simple or Oriental ...

How to Tie a Tie: The Complete Guide | The Art of Manliness
https://www.artofmanliness.com/articles/how-to-tie-a-tie/
May 10, 2019 - It's a sad fact, but there are grown men who don't know **how to tie a tie**. If they have a big interview that afternoon, they'll go shopping for a ...

How To Tie A Tie | Tie Knots - Brooks Brothers
https://www.brooksbrothers.com/how-to-tie-a-tie/how-to-tie-a-tie,default,pg.html
Watch our **how to tie a tie** videos on four classic knots including the Bow Tie knot, Windsor knot, Half Windsor knot, and Four in Hand knot. The four tie knots ...

Popular Ways to Tie a Necktie | Bows-N-Ties.com
https://www.bows-n-ties.com/how-to-tie-a-necktie/
The **tie** knots listed below are sorted by difficulty, starting with the popular and easy to master Four-in-Hand knot, and ending with the more challenging dimpled ...

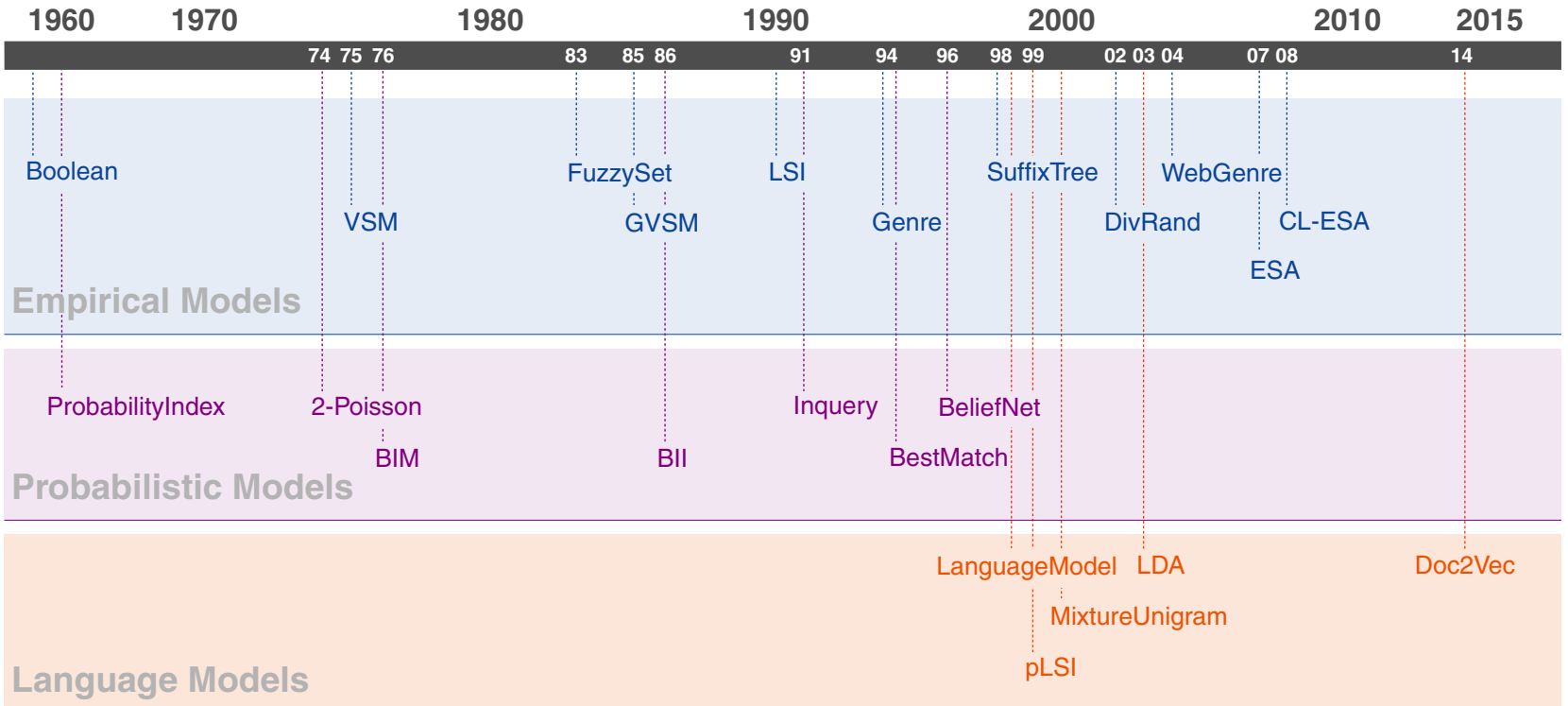
Tie-a-Tie.net
https://www.tie-a-tie.net/
Learn **how to tie a tie** with the Windsor, Half Windsor, Four in Hand and Pratt necktie knots by following step-by-step video instructions and colored diagrams ...

How to Tie a Half Windsor Knot | The Tie Bar
https://www.thetiebar.com/how-to-tie/half-windsor
Learn **how to tie** a Half Windsor Knot and quickly become a Half Windsor expert with our easy, step-by-step instructions.

©Völske 2019

Enhancing Result Rankings with Axioms

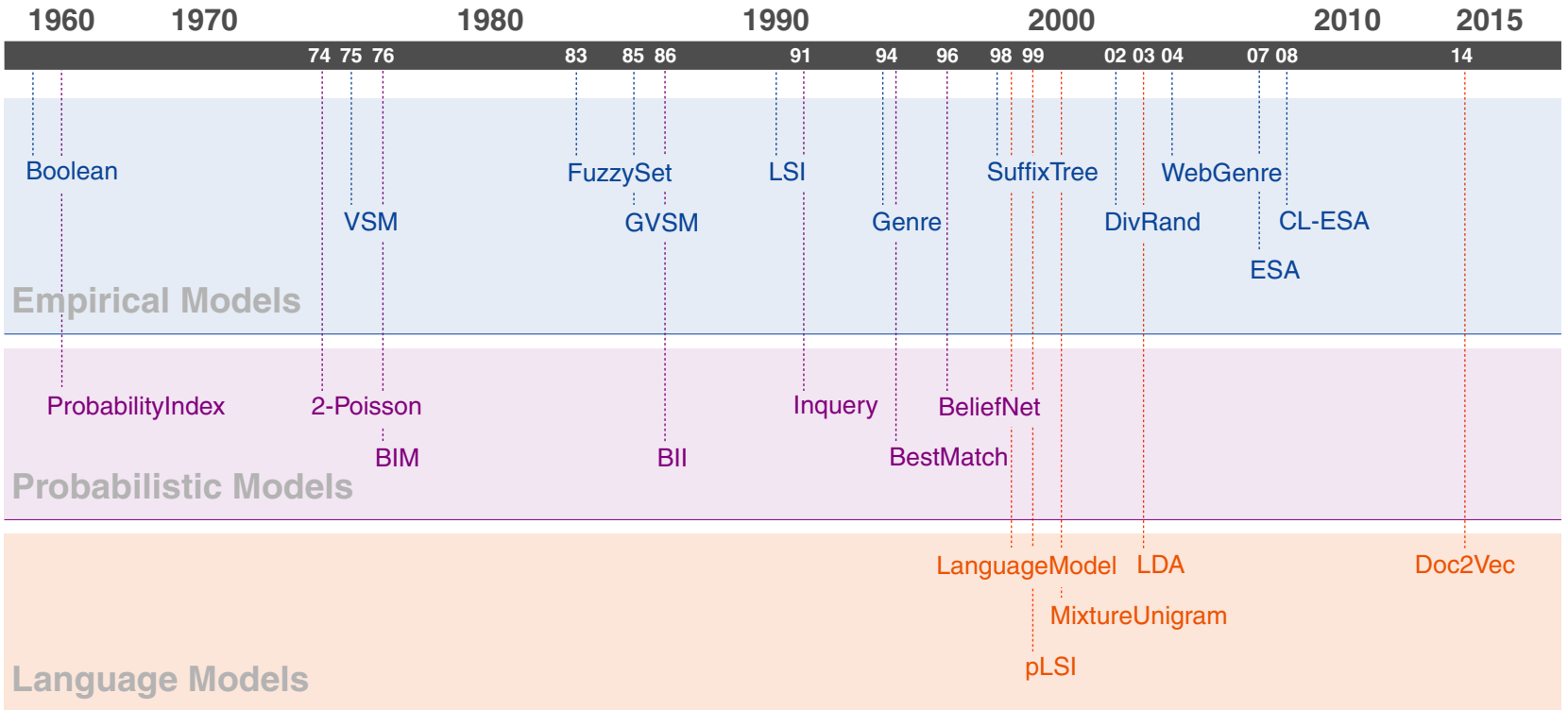
Retrieval Models



© Völske 2019

Enhancing Result Rankings with Axioms

Retrieval Models



Strong retrieval models share similar heuristics

© Völske 2019

Enhancing Result Rankings with Axioms

Retrieval Axioms

Purpose	Axioms	Source
Term frequency	TFC1–TFC3	[Fang, Tao, Zhai; SIGIR'04]
	TDC	[Fang, Tao, Zhai; SIGIR'04]
Document length	LNC1 + LNC2	[Fang, Tao, Zhai; SIGIR'04]
	TF-LNC	[Fang, Tao, Zhai; SIGIR'04]
	QLNC	[Cummins, O'Riordan; CIKM'12]
Lower bound	LB1 + LB2	[Lv, Zhai; CIKM'11]
Query aspects	REG	[Zheng, Fang; ECIR'10]
	DIV	[Gollapurdi, Sharma; WWW'09]
Semantic similarity	STMC1 + STMC2	[Fang, Zhai; SIGIR'06]
	STMC3	[Fang, Zhai; SIGIR'06]
	TSSC1 + TSSC2	[Fang, Zhai; SIGIR'06]
Term proximity	PHC + CCC	[Tao, Zhai; SIGIR'07]

©Völske 2019

Enhancing Result Rankings with Axioms

Retrieval Axioms

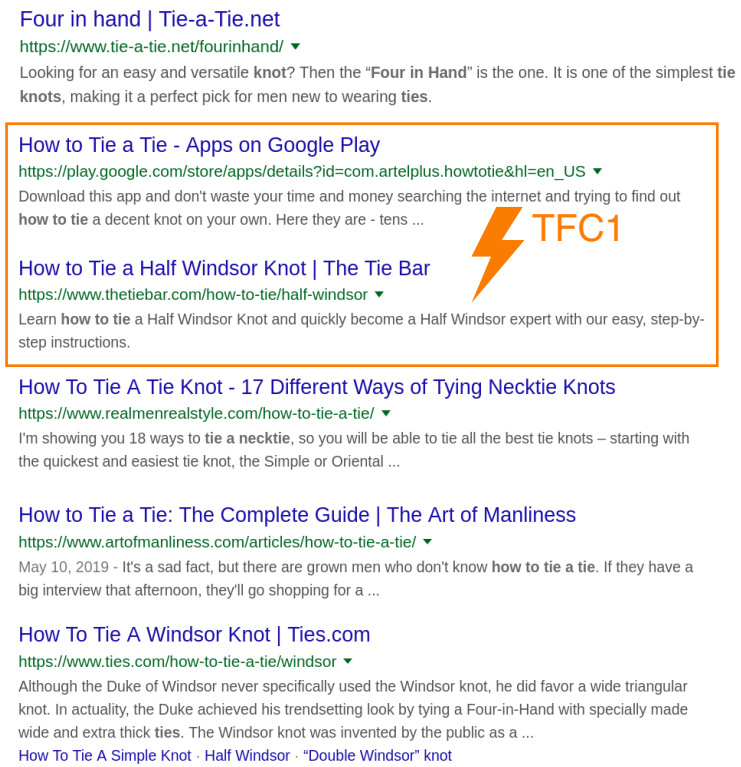
Purpose	Axioms	Source
Term frequency	TFC1–TFC3	[Fang, Tao, Zhai; SIGIR'04]
	TDC	[Fang, Tao, Zhai; SIGIR'04]
Document length	LNC1 + LNC2	[Fang, Tao, Zhai; SIGIR'04]
	TF-LNC	[Fang, Tao, Zhai; SIGIR'04]
	QLNC	[Cummins, O'Riordan; CIKM'12]
Lower bound	LB1 + LB2	[Lv, Zhai; CIKM'11]
Query aspects	REG	[Zheng, Fang; ECIR'10]
	DIV	[Gollapurdi, Sharma; WWW'09]
Semantic similarity	STMC1 + STMC2	[Fang, Zhai; SIGIR'06]
	STMC3	[Fang, Zhai; SIGIR'06]
	TSSC1 + TSSC2	[Fang, Zhai; SIGIR'06]
Term proximity	PHC + CCC	[Tao, Zhai; SIGIR'07]

Only used in theoretical analysis of retrieval models so far.

© Völske 2019

Axiomatic Result Re-ranking

Research Questions



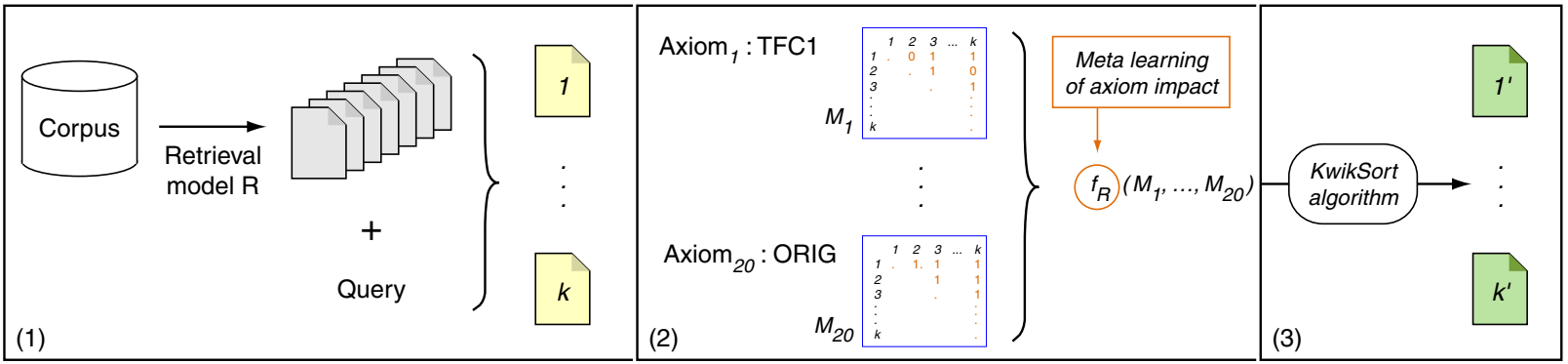
Can fixing axiom violations improve individual rankings?

How to incorporate axioms directly into the retrieval process?

© Völske 2019

Axiomatic Result Re-ranking

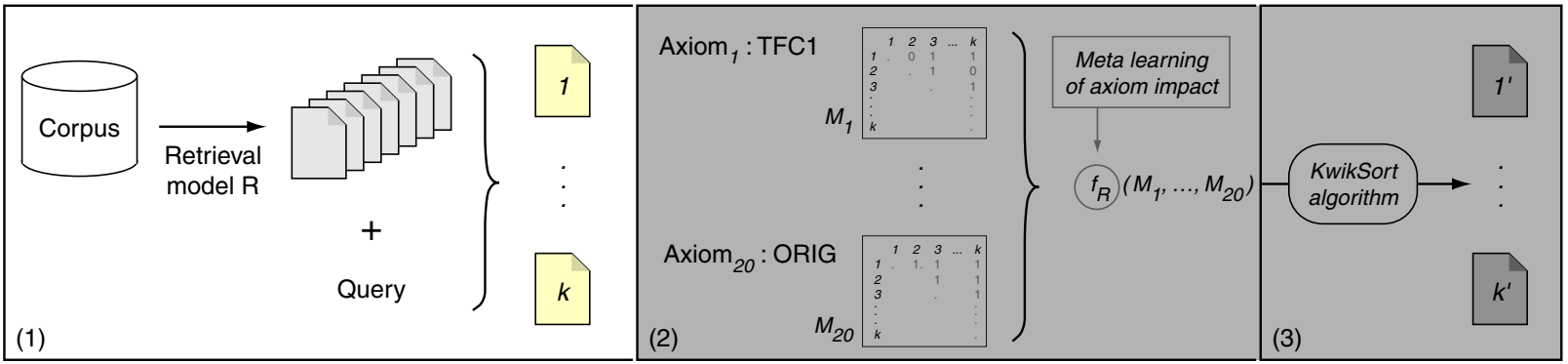
Axiomatic Re-ranking Pipeline



© Völske 2019

Axiomatic Result Re-ranking

Axiomatic Re-ranking Pipeline

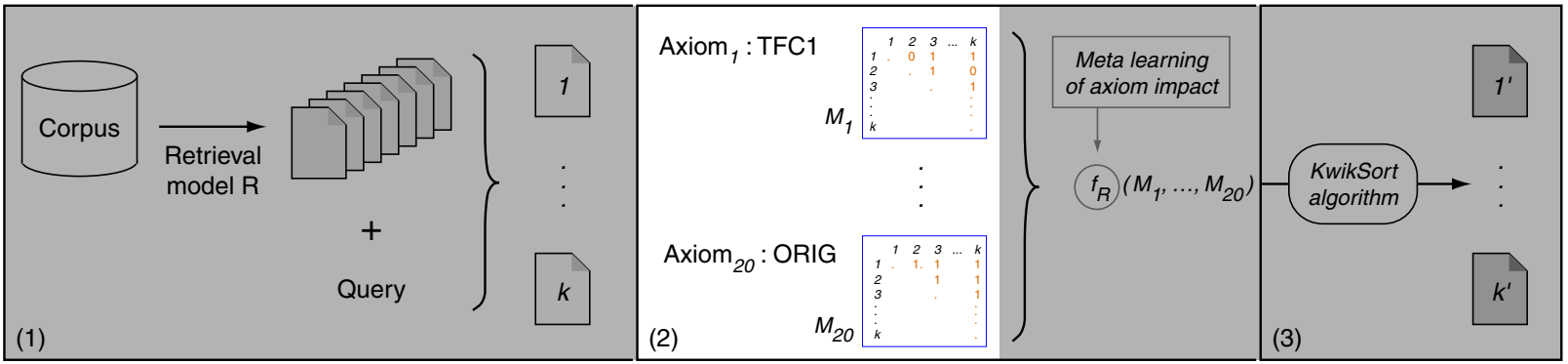


1. Retrieve an initial top-k result set.

© Völske 2019

Axiomatic Result Re-ranking

Axiomatic Re-ranking Pipeline



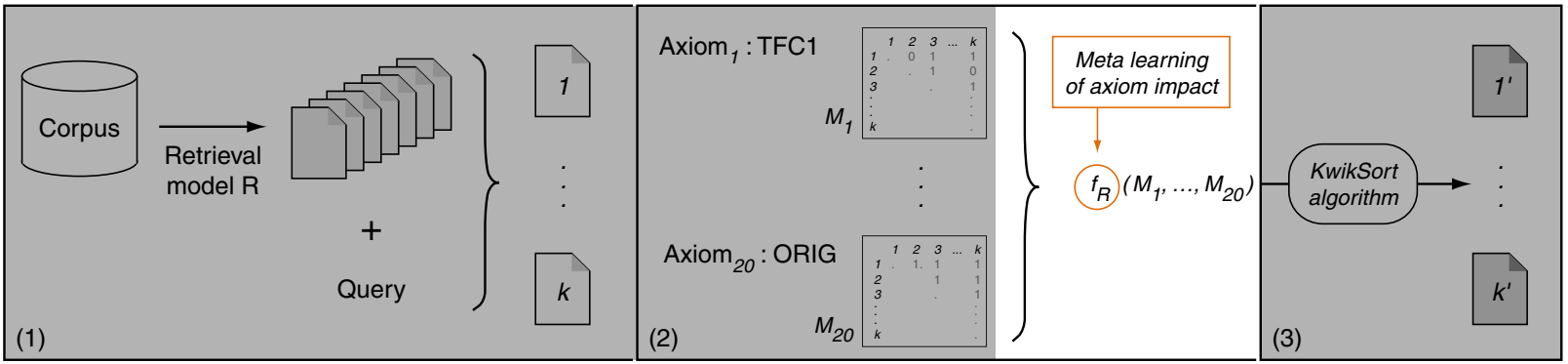
1. Retrieve an initial top-k result set.

2. (a) Compute re-ranking preferences of various axioms.

©Völske 2019

Axiomatic Result Re-ranking

Axiomatic Re-ranking Pipeline



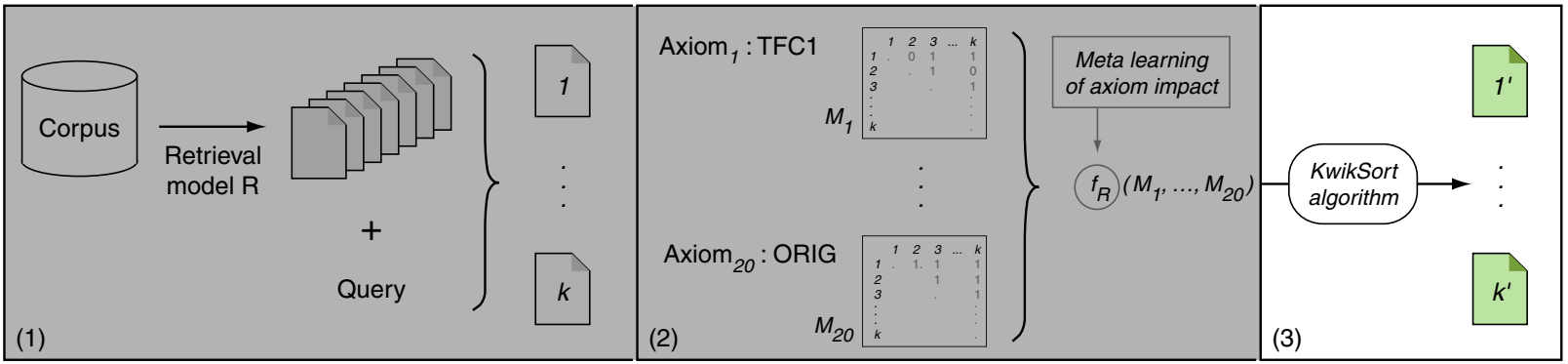
1. Retrieve an initial top-k result set.

2. (a) Compute re-ranking preferences of various axioms.
   (b) Aggregate re-ranking preferences.

© Völske 2019

Axiomatic Result Re-ranking

Axiomatic Re-ranking Pipeline

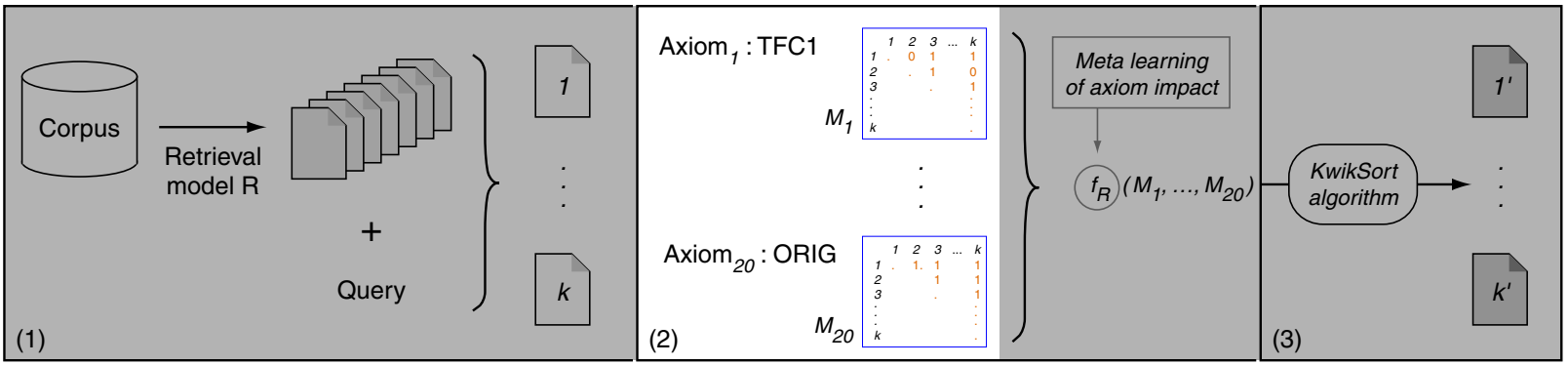


1. Retrieve an initial top-k result set.
2. (a) Compute re-ranking preferences of various axioms.
   (b) Aggregate re-ranking preferences.
3. Re-rank the initial result set.

© Völske 2019

Axiomatic Result Re-ranking

Detail Step 2a

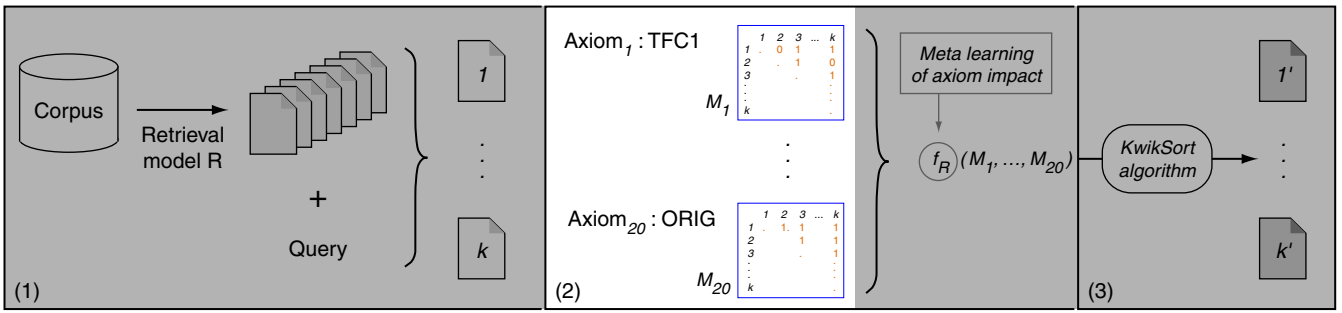


1. Retrieve an initial top-k result set.
2. (a) Compute re-ranking preferences of various axioms.
   (b) Aggregate re-ranking preferences.
3. Re-rank the initial result set.

© Völske 2019

Axiomatic Result Re-ranking

Detail Step 2a



Re-state existing axioms as binary predicates:

$$A(q, d_1, d_2) = \begin{cases} 1 & \text{if } d_1 >_A d_2 \text{ given } q \\ 0 & \text{otherwise.} \end{cases}$$

q	query
d_1, d_2	documents
$d_1 >_A d_2$	d_1 should be ranked higher according to axiom A

© Völske 2019

Axiomatic Result Re-ranking

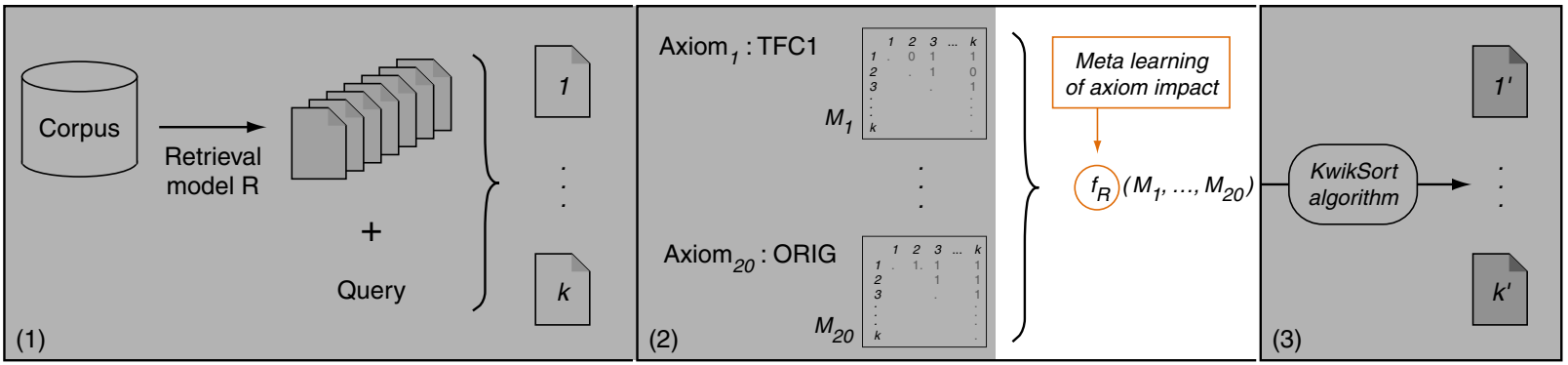
Adapting Existing Axioms to Express Re-Ranking Preferences

Purpose	Acronyms	Adapted
Term frequency	TFC1–TFC3	✓
	TDC	✓
Document length	LNC1 + LNC2	✓
	TF-LNC	✓
	QLNC	✗
Lower bound	LB1 + LB2	✓
Query aspects	REG	✓
	DIV	✓
Semantic similarity	STMC1 + STMC2	✓
	STMC3	✗
	TSSC1 + TSSC2	✗
Term proximity	PHC + CCC	✗
	QPHRA	New
	PROX1–5	New
Other	ORIG	New

©Völske 2019

Axiomatic Result Re-ranking

Detail Step 2b

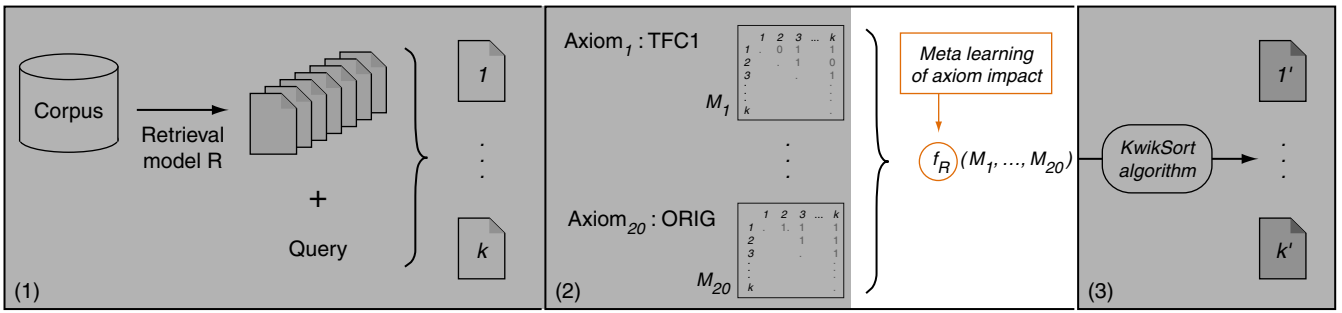


1. Retrieve an initial top-k result set.
2. (a) Compute re-ranking preferences of various axioms.
   (b) Aggregate re-ranking preferences.
3. Re-rank the initial result set.

© Völske 2019

Axiomatic Result Re-ranking

Detail Step 2b



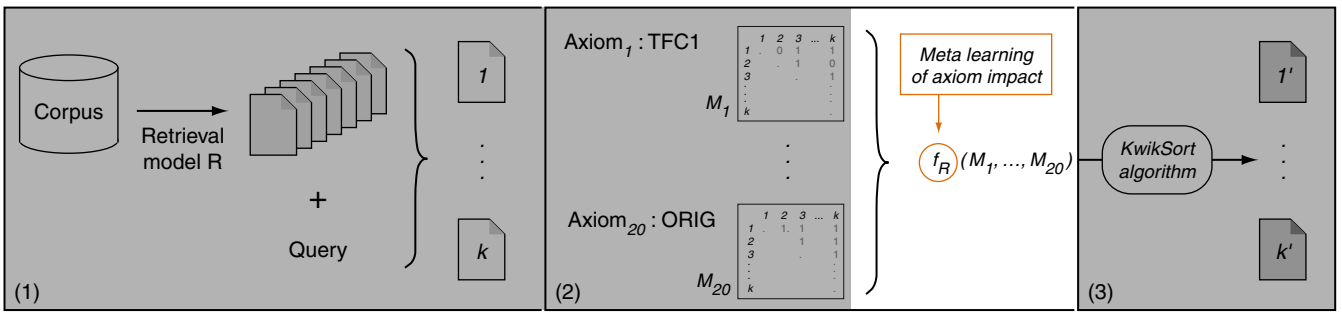
Given a top-k ranking, each axiom produces a k-by-k preference matrix.

Goal: Aggregate the individual preference matrices for a set of axioms.

© Völske 2019

Axiomatic Result Re-ranking

Detail Step 2b



Given a top-k ranking, each axiom produces a k-by-k preference matrix.

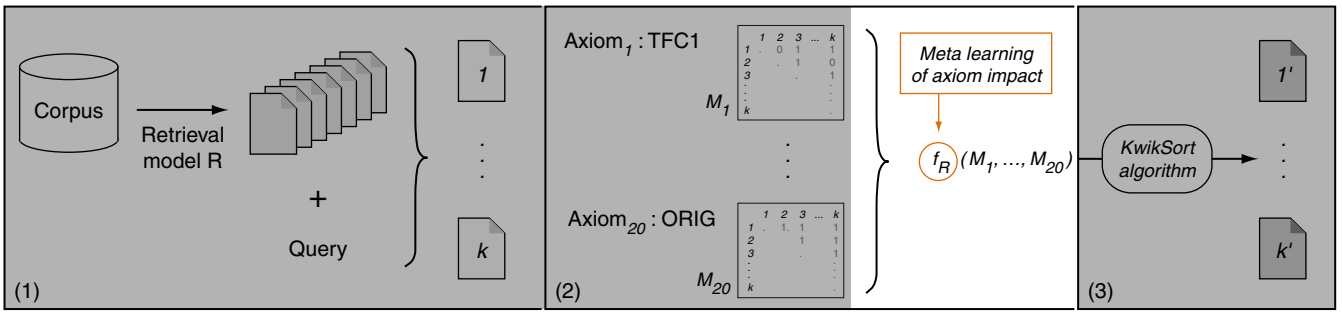
Goal: Aggregate the individual preference matrices for a set of axioms.

Hypothesis: Rankings from different retrieval models deviate from the axiomatic constraints in different ways.

© Völske 2019

Axiomatic Result Re-ranking

Detail Step 2b



Given a top-k ranking, each axiom produces a k-by-k preference matrix.

Goal: Aggregate the individual preference matrices for a set of axioms.

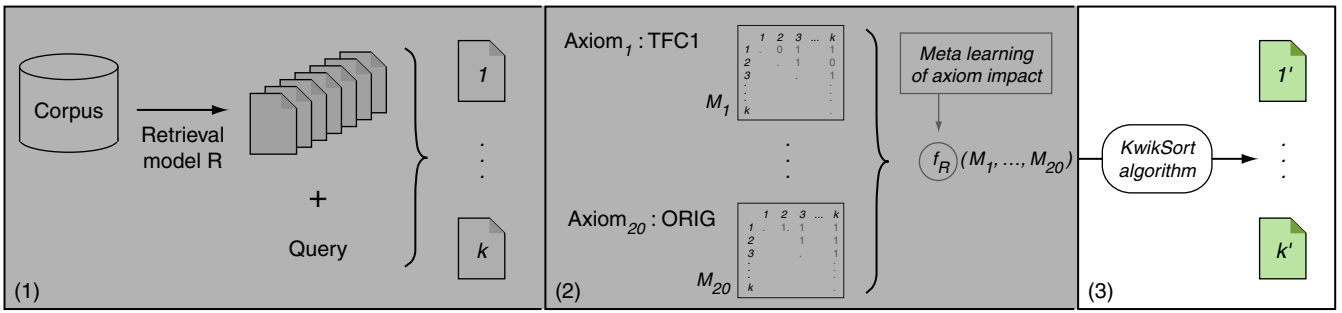
Hypothesis: Rankings from different retrieval models deviate from the axiomatic constraints in different ways.

Approach: Learn a preference aggregation function specific to the retrieval model.

© Völske 2019

Axiomatic Result Re-ranking

Detail Step 3



Aggregated preference matrix may still contain contradictions.

Resolve via minimum feedback arc set approximation.

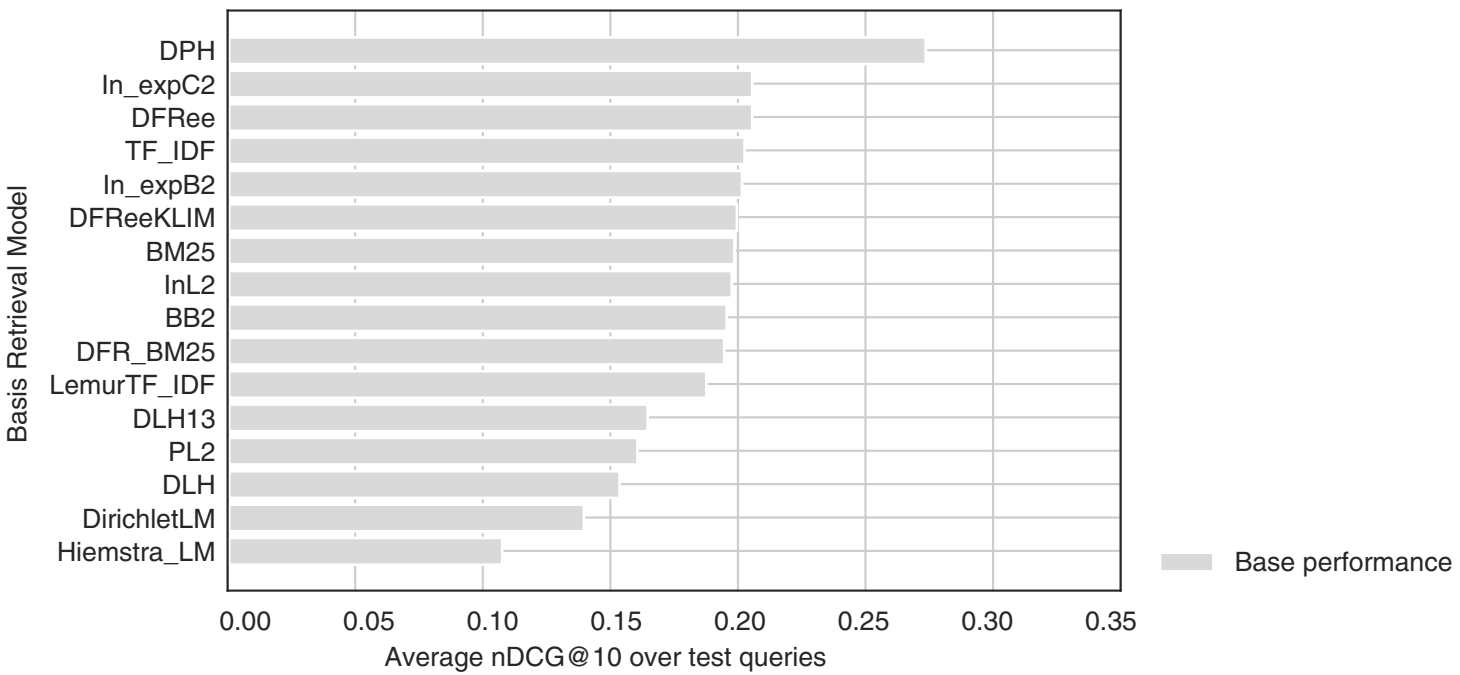
© Völske 2019

Experimental Evaluation

Fixing axiom violations in individual rankings does help.

16 different retrieval models over ClueWeb09

120 training queries + 60 test queries

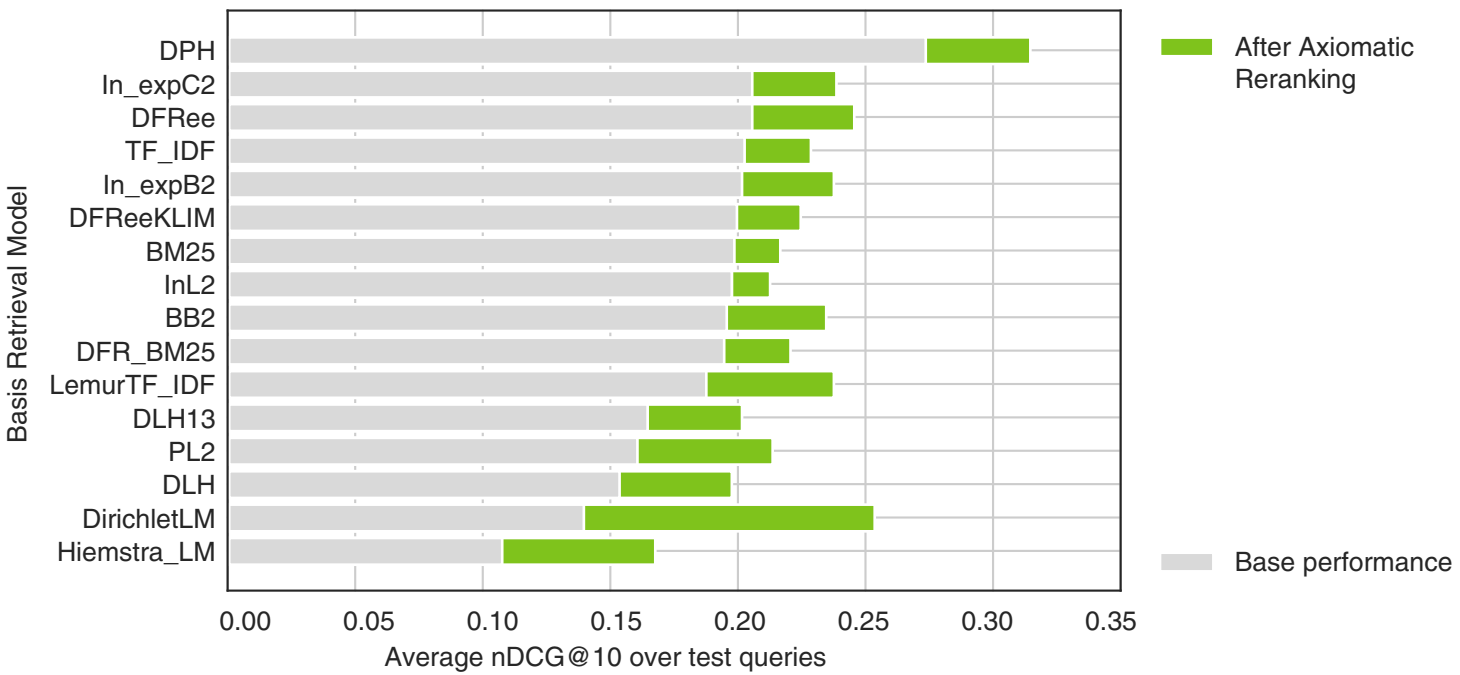


© Völske 2019

Experimental Evaluation

Fixing axiom violations in individual rankings does help.

16 different retrieval models over ClueWeb09
120 training queries + 60 test queries



© Völske 2019

Summary

Authors using a search engine. . .

- Tend follow one of two opposite strategies
- Can be supported by search engines that predict retrieval success

Search engine result rankings. . .

- Can be made to conform to theoretically sound retrieval axioms
- Especially if the basis retrieval model is weak

©Völske 2019

Benno Stein • Pertti Vakkari • Norbert Siegmund • Florian Echtler • Andreas Jakoby • Magdalena Wolska • Martin Potthast • Matthias Hagen • Henning Wachsmuth • Tim Gollub • Johannes Kiesel • Shahbaz Syed • Wei-Fan Chen • Roxanne El-Baff • Khalid Alkhatib • Yamen Ajjour • Janek Bevendorff • Matti Wiegmann • Alexander Bondarenko • Milad Ashomary • Kevin Lang • Nathalie Dittrich • Maik Anderka • Steven Burrows • Nedim Lipka • Dennis Hoppe • Sven Meyer zu Eißen • Peter Prettenhofer • Martin Trenkmann • Edgardo Ferretti • Elisabeth Lex • Marcelo Errecalde • Leticia Cagnina • Christopher Horn • Michael Granitzer • Pavel Braslavski • Galina Lezina • Jörg Frochte • Ana Belén Martínez Torres • Steve Göring • Christof Bräutigam • Christian Forler • Eik List • Jakob Wenzel • Nadin Glaser • Melanie Henning • Maria-Theresa Hansens • Tina Meinhardt • Christin Oehmichen • Susanne Robbel • Anja Ludovici • Martin Tippmann • Martin Heinrich • Alexander Kümmel • Alexander Herr • Maximilian Michel • Christian Fricke • Jan Grassegger • Kai Lorenz • Jakob Gomoll • Sebastian Hönel • Abdul Hamid Sabri • Marie Bornemann • Lene Ganschow • Florian Kneist • Goran Košutić • Manuel Willem • Kristof Komlossy • Jonas Köhler • André Karge • Benedikt Vogler • Bagrat Ter-Akopyan • Olaoluwa Phillip Anifowose • Oana-Iuliana Popescu • Erika Garces • Saeed Entezari • Rosemary Adejoh • Katarzyna Jabłonska • Christian Becker • Matthew Heinz • Erdan Genc • Payam Adineh • Ehsan Fatehifar • Jiani Qu • Tristan Licht • Josef Roth • Sybille, Jörg und Martin Völske • Jasmin Türker

Colleagues, Collaborators, Students, Friends & Family...

Thank you!